Exhibit 20

Human Molecular Genetics, 1997, Vol. 6, No. 8 1241–1250

Autism and multiple exostoses associated with an X;8 translocation occurring within the *GRPR* gene and 3′ to the *SDC2* gene

Yumiko Ishikawa-Brush[1], John F. Powell[2], Patrick Bolton[3], Andrew P. Miller[4,+], Fiona Francis[5], Huntington F. Willard[4], Hans Lehrach[5] and Anthony P. Monaco[1,*]

[1]Wellcome Trust Centre for Human Genetics, University of Oxford, Windmill Road, Headington, Oxford OX3 7BN, UK, [2]Department of Neuroscience, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF, UK, [3]University of Cambridge, Section of Developmental Psychiatry, 18b Trumpington Road, Cambridge CB2 2AH, UK, [4]Department of Genetics and Center for Human Genetics, Case Western Reserve University, Cleveland, OH 44106, USA and [5]Max-Planck-Institut für Molekulare Genetik Ihnestrasse 73, 14195 Berlin, Germany

Received March 3, 1997; Revised and Accepted May 13, 1997

An X;8 translocation was identified in a 27-year-old female patient manifesting multiple exostoses and autism accompanied by mental retardation and epilepsy. Through molecular analysis using yeast artificial chromosomes (YACs) and cosmid clones, the translocation breakpoint was isolated and confirmed to be reciprocal within a 5′-GGCA-3′ sequence found on both X and 8 chromosomes without gain or loss of a single nucleotide. The translocation breakpoint on the X chromosome occurred in the first intron of the gastrin-releasing peptide receptor (*GRPR*) gene and that on chromosome 8 occurred ~30 kb distal to the 3′ end of the Syndecan-2 gene (*SDC2*), also known as human heparan sulfate proteoglycan or fibroglycan. The *GRPR* gene was shown to escape X-inactivation. A dosage effect of the *GRPR* and a position effect of the *SDC2* gene may, however, contribute the phenotype observed in this patient since the orientation of these genes with respect to the translocation was incompatible with the formation of a fusion gene. Investigation of mutations in these two genes in unrelated patients with either autism or multiple exostoses as well as linkage and association studies is needed to validate them as candidate genes.

INTRODUCTION

Rare X;8 autosomal translocations in females have been very important in identifying genes directly involved in disease phenotypes (1). A 27-year-old female patient (ML) with a balanced translocation 46,X,t(X;8)(p22.13;q22.1) has been described with multiple exostoses and autism. Details of her clinical condition and cytogenetic analysis are described in Bolton *et al.* (2). Briefly, multiple exostoses were present around ankles, knees, wrists and left clavicle, and the patient had short stature, short span and small hands with short 4th and 5th metacarpals. She has a mild brachycephaly and was diagnosed as autistic according to the International Classification of Diseases Criteria following evaluations using the Autism Diagnostic Interview (3) and the Autism Diagnostic Observation Schedule (4). She is mentally retarded with an IQ of 35 and also suffers from grand mal epilepsy. The translocated X chromosome of the patient, ML, was demonstrated to be active by bromodeoxyuridine analysis (2). The chromosomal translocation was not seen in either her mother or brother, and their karyotypes are normal. They had no clinical signs of multiple exostoses or autism. Her father is deceased and his karyotype was not determined.

To understand further the phenotype in this patient, positional cloning was undertaken to isolate the translocation breakpoint and search for genes possibly affected by the rearrangement.

Autism (MIM 209850) (5) is a neurodevelopmental disorder characterized by qualitative impairments in communication and reciprocal social interaction as well as restricted, repetitive, stereotyped patterns of behaviour, interests and activities. It frequently is accompanied by mental retardation (75%) and epilepsy (25%) (6,7). The incidence in the general population is estimated to be ~2–4 per 10 000 children. The neurobiological defect is unknown, but autism has been shown by both twin and family studies to have a strong genetic aetiology (2–8). Fragile X is found in ~5% of cases and tuberous sclerosis in 1–2% (13,14), indicating some genetic heterogeneity.

The sex ratio of autism is 3–4 times higher in males than in females (7). This suggests a susceptibility locus for autism may be X-linked, or the phenotype is partially sex-limited. However, a major X-linked locus recently has been ruled out by segregation and linkage analysis of multiplex autism families (15,16). Case reports have shown links between autism and gross chromosomal abnormalities, including monosomy and trisomy of large chromosome regions (17–24). In one of the above case reports,

*To whom correspondence should be addressed. Tel: +44 1865 740019; Fax: +44 1865 742186; Email: anthony.monaco@well.ox.ac.uk
+Present address: Sequana Therapeutics Inc., 11099 North Torrey, Pines Road, Suite 160, La Jolla, CA 92037, USA

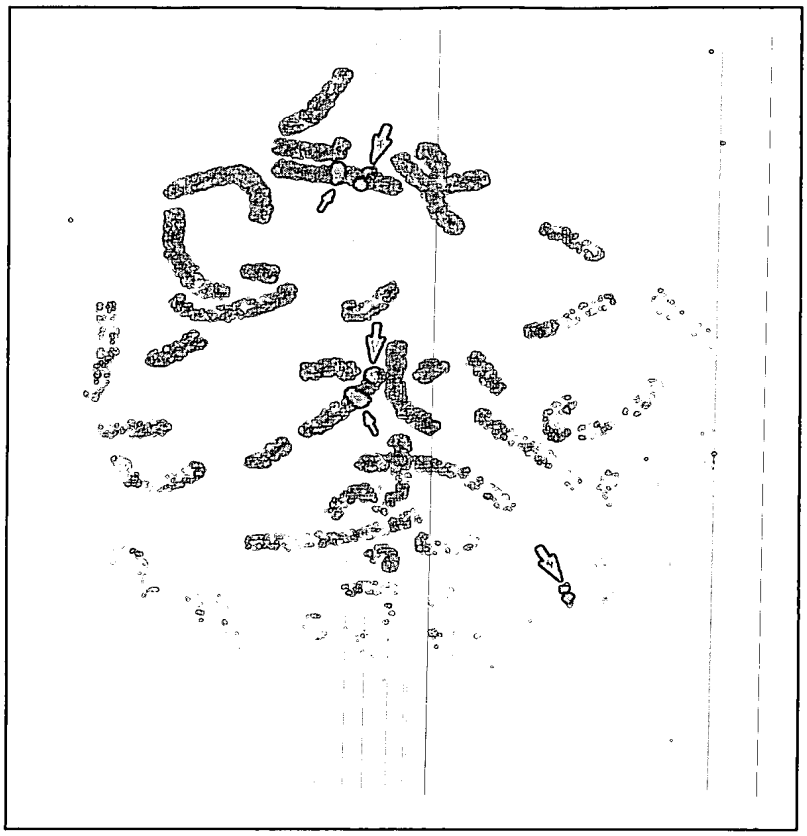


Figure 1. Hybridizations of YAC-57 on normal X, der(X) and der(8) chromosomes (large arrow) indicate that the YAC spans the translocation breakpoint. The X alpha satellite marker was used to identify the X chromosome centromere (small arrow).

one autistic patient had a duplication of Xp22.1–p22.3 (24), covering the same region as ML's translocation breakpoint.

Multiple exostoses (EXT: MIM 133700) (5) are characterized by cartilaginous protuberances at the ends of the diaphyses of the extremities. Exostoses are rarely present at birth, but gradually develop from the growth plate between the epiphysis and metaphysis. They increase in size and undergo ossification, giving rise to skeletal deformities and short stature (25). The frequency of multiple exostoses has been estimated to be 1 in 50 000 to 1 in 100 000 in the West (25,26).

EXT is genetically heterogeneous, with three loci (*EXT1*, *EXT2* and *EXT3*) mapped to 8q24.1, 11p11 and 19p, respectively. The *EXT1* and *EXT2* genes have been cloned and sequenced (27–29). Both genes are significantly homologous to each other, suggesting derivation from a common ancestral gene that gave rise to an *EXT* multigene family (29). The functions of the *EXT* genes are unknown, but sequence and mutation analyses suggest they may act as tumour-suppressor genes (29,30).

Further heterogeneity of EXT may occur on chromosome 8. Yoshiura *et al.* reported an 8q23 translocation in association with multiple exostoses (31), and the patient, ML, has a translocation in the region of 8q22.1 (2).

RESULTS

Isolation of YACs spanning the translocation breakpoint

To determine flanking markers and identify yeast artificial chromosomes (YACs) spanning the translocation breakpoint in Xp22, fluorescence *in situ* hybridization (FISH) was applied to YACs isolated from the ICRF human YAC library (32). This library has been screened using X-chromosome probes (*POLA*, *ZFX*, *DXS41*, *DXS274*, *DXS257*, *DXS365*, *DXS43*, *DXS197*, *DXS207*, *GLYR* and *DXS9*) covering the region from Xp22.11 to p22.31 (33,34). Using FISH, 42 YACs were examined of which 23 were proximal to the breakpoint, 14 were distal, and five crossed the translocation breakpoint. The proximal and distal flanking markers were *DXS43* and *DXS197*, respectively. About 28% of the YACs from this region were found to be chimeric and 10% contained two YACs. The five YACs crossing the translocation breakpoint were YAC-57 (ICRFy900B0544), YAC-58 (ICRFy900E1092), YAC-60 (ICRFy900D0758), YAC-61 (ICRFy900E1112) and YAC-63 (ICRFy900B1148) (Fig. 1).

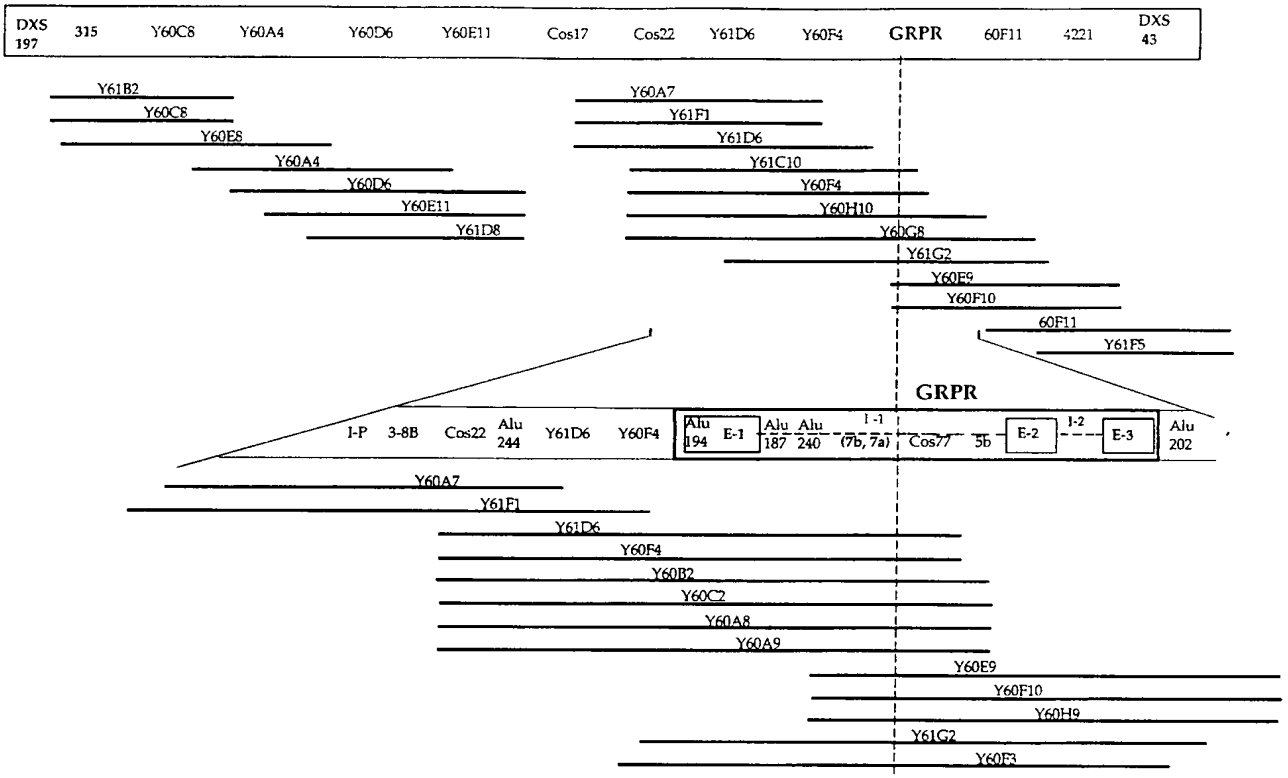


Figure 2. Cosmid contig and markers around Xp22. The *GRPR* region is magnified. Probes including STSs, Alu-PCR and whole cosmids are indicated above the contig. One cosmid is ~30–40 kb.

Construction of the cosmid contig and localization of the *GRPR* gene

The two shortest YACs, YAC-60 (480 kb) and YAC-61 (520 kb), were used to isolate cosmids by whole YAC screening of the ICRF flow-sorted X cosmid library and to construct cosmid sublibraries. A total of 250 cosmids were isolated and ordered in a contig using the probes generated by cosmid–cosmid and Alu-PCR hybridization methods, and also probes derived from the *GRPR* gene sequence. The contig covered 500 kb with one gap (Fig. 2).

The *GRPR* cDNA sequence of ~1.6 kb is contained in three exons (35). Primers were designed to amplify by PCR exons 1–3, and introns 1 and 2. Intron 2 was ~1.5 kb in length, but intron 1 appeared to be too long for standard PCR amplification. The PCR products were used as hybridization probes to localize the *GRPR* gene in the cosmid contig. Eight cosmids were found to cover the whole *GRPR* gene, which indicated the *GRPR* gene to be up to 40 kb in length (Fig. 2).

Identification of the junction fragment within the *GRPR* gene

Three of the eight cosmids covering the *GRPR* region were selected for hybridization to blots containing genomic DNA from ML digested with *Eco*RI, *Eco*RV, *Pst*I and *Bgl*II. The cosmids identified obvious junction fragments on all enzyme digests (data not shown). Two single copy *Hin*dIII subclones (7a and 7b), derived from cosmid Y61G2, also identified junction fragments that were unique to each probe, indicating that they may flank the breakpoint on the X chromosome (7b shown in Fig. 3). Both subclones hybridized to cosmids Y60E9, Y60F10 and Y60H9. These contain exon 2, intron 2 and exon 3, but not exon 1 of the *GRPR* gene. The subclones also hybridized with cosmids Y60F4, Y60B2, Y60C2, Y60A8 and Y60A9, which are positive for exon 1 of the *GRPR* gene, but not exon 2, intron 2 nor exon 3. These results placed subclones 7a and 7b within the first intron of the *GRPR* gene, proximal to exon 1 (Fig. 2).

Cloning the junction fragment and construction of a restriction map

Genomic DNA from ML was digested with *Bgl*II and a phage library was constructed and screened with the *Hin*dIII subclones 7a and 7b. Positive phage clones were isolated and subcloned into plasmids. Three types of plasmid clones were identified: (i) positive with both 7a and 7b, containing a 4.6 kb *Bgl*II fragment derived from the normal X chromosome; (ii) strongly positive with 7a only, and containing a 5.0 kb *Bgl*II fragment; and (iii) positive with 7b, and containing a 5.5 kb *Bgl*II fragment. The latter two *Bgl*II fragments are the two separate translocation junction fragments derived from either the der(X) or der(8)

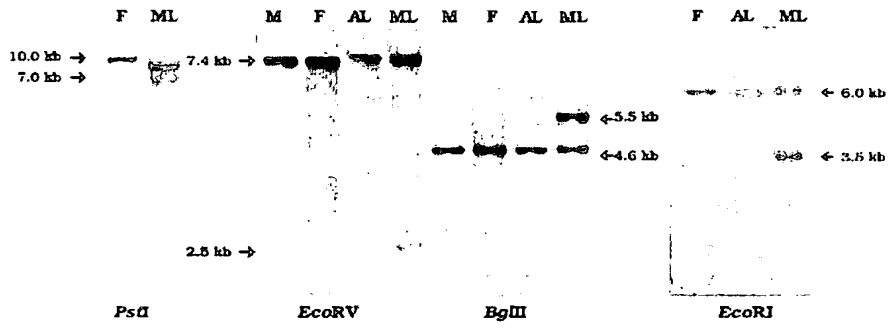


Figure 3. Probe 7b detection of the translocation junction fragment in patient genomic DNA digested with *Eco*RI, *Bgl*II, *Eco*RV and *Pst*I. Key: ML, patient; AL, patient's normal mother; F, female; M, male.

chromosomes. A restriction map of these three subclones was constructed using single, double and triple enzyme digestions and hybridization with 7a and 7b. The restriction map (Fig. 4A) confirmed the translocation breakpoint to be within subclone 7a. Long PCR amplification between the junction fragment and exon 1 and 2, indicated that the *Bgl*II 5.5 kb junction fragment was located on the der(8) and the *Bgl*II 5.0 kb fragment on the der(X). The orientation of the junction fragments with respect to the centromere and telomere was determined by sequencing and is also shown in Figure 4A.

Identification of a DNA sequence spanning the X;8 translocation breakpoint

Inserts of these normal X, der(X) and der(8) subclones and the 7a and 7b subclones were partially sequenced to identify the translocation breakpoint region. The combined sequence data showed: (i) 7b is adjacent to 7a, and (ii) the translocation breakpoint is located within 7a, 200 bp away from 7b, in the middle of primer-388 (Fig. 4A). This result is consistent with hybridization and PCR analyses (data not shown).

To examine the mechanism of the translocation, the regions of the breakpoint from the two normal and two derived translocation chromosomes were sequenced and compared. A motif, GGCA, was found on all four chromosomes. The translocation occurred within the 5′-GGCA-3′ sequence motif without the loss or gain of a single nucleotide (Fig. 4B).

Expression of *GRPR*

In previous studies on X;autosome translocations, affected females were nullisomic for the responsible locus, as the X-linked gene on the normal X was subject to X inactivation and the translocation disrupted the copy on the active X. To determine whether *GRPR* was subject to inactivation, we assayed its expression from normal active and inactive X chromosomes, using RT-PCR of RNA isolated from a panel of mouse–human somatic cell hybrids. The panel consisted of two active X-containing hybrids, Xa1 (AHA-11aB1) and Xa2 (t60-12), and four inactive X-containing hybrids, Xi1 (t11-4Aaz5), Xi2 (t48-1a-1Daz4a), Xi3 m(t75-2maz34-4a) and Xi4 (t86-B1maz1b-3a) (36–38). Expression of *GRPR* was detected from both the active and inactive X chromosomes, indicating that it escapes X-inactivation (Fig. 5).

On Northern blot analysis of poly(A)$^+$ RNA (Clontech), *GRPR* expression was detected in pancreas, but not in brain, placenta, kidney, prostate, testis, small intestine, colon, heart, lung, liver, skeletal muscles and ovary (data not shown). Northern blot hybridization to total RNA isolated from control and patient lymphoblastoid cells was also negative.

Identification of the translocation breakpoint on chromosome 8

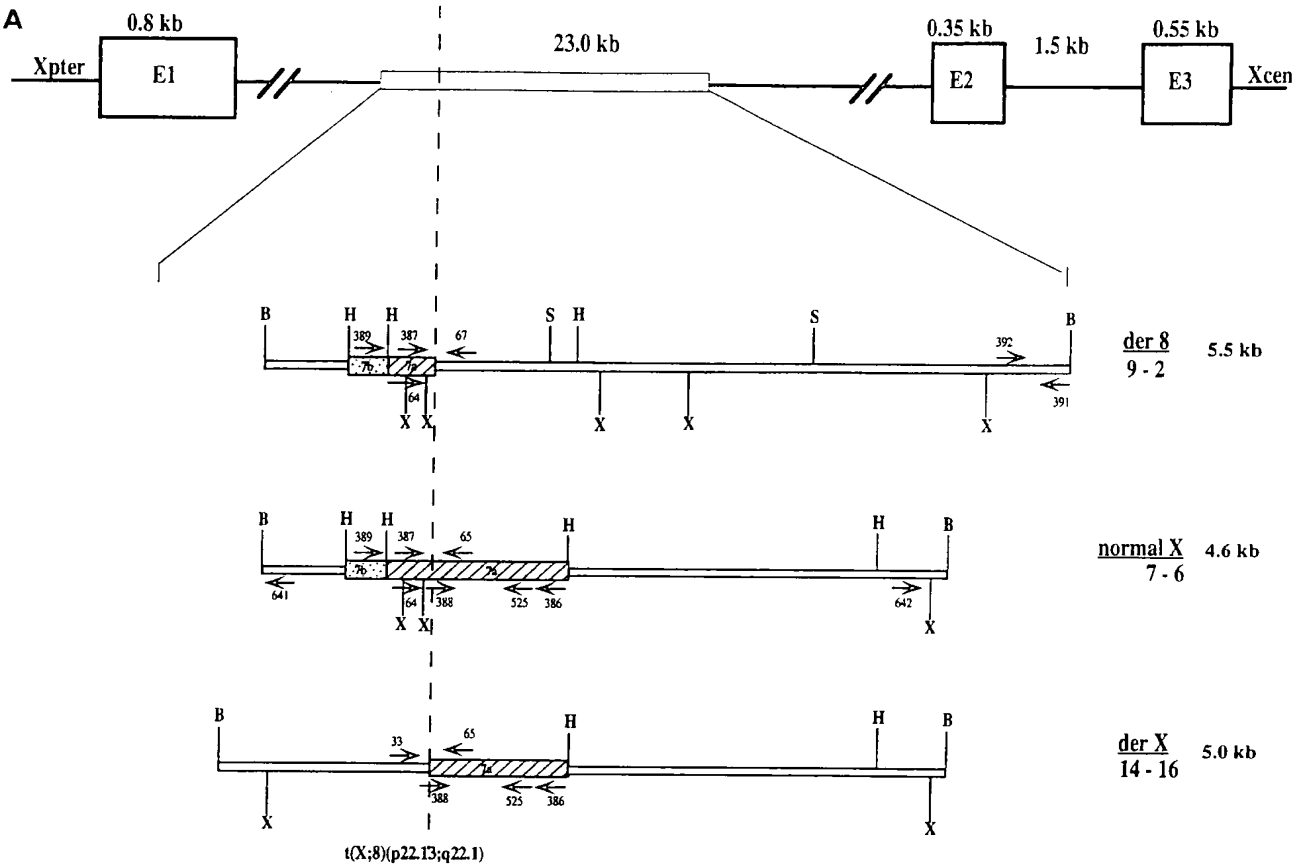
Chromosome 8 was analysed by sequencing across the translocation breakpoint using primers from both the der(X) and der(8) plasmids. A 240 bp PCR product, I33-67, produced by these primers was confirmed to be chromosome 8 specific by amplifying DNA from a somatic cell hybrid specific for chromosome 8 but not for chromosome X (Fig. 4A). On Southern blot analysis of patient ML, I33-67 also identified translocation junction fragments with several enzymes, confirming that it was adjacent to the breakpoint (data not shown).

To search for a gene on chromosome 8 around this region, I33-67 was used to isolate three YACs, 2B-B6 (550 kb), 13B-G12 (500 kb) and 38B-F2 (700 kb), by PCR screening of the ICI YAC library. Cosmid libraries were constructed from two of the shortest and non-chimeric YACs (2B-B6 and 13B-G12) and were screened by hybridization using I33-67 as a probe.

Identification of the Syndecan-2 (*SDC2*) gene

Two cosmids, C2E and C7B (Fig. 6), out of seven cosmids spanning the translocation breakpoint, were used as probes by hybridization to screen a human testis cDNA library. Several cDNAs were purified, subcloned and sequenced. The BLASTN search identified the SDC2 gene with 100% homology from each of these cDNA fragments. This gene is also known as human heparan sulfate proteoglycan or fibroglycan (MIM: J04621) (5) and previously had been localized by Marynen *et al.* to 8q22–q24 (39).

A cosmid contig covering this region was constructed using end fragments subcloned from cosmids (Fig. 6). The *SDC2* gene sequence in the database covers 3.4 kb of the 3′ end, and the three YACs and cDNAs isolated with I33-67 contained four exons also from the 3′ end. The contig confirmed that the translocation breakpoint did not disrupt the *SDC2* gene, but occurred <30 kb distal to the 3′ end (Fig. 6). FISH analysis of cosmids C10D, C6E, C1D and C1G confirmed the orientation of the *SDC2* gene with respect to the translocation breakpoint, and demonstrated that a



B

```
NX: 7-6   GGTAATTAGGAAGTGGCATTACATCTGATGCAGT
                  |||||||||||||||||
9-2       GGTAATTAGGAAGTGGCAgcagaattggagttca
                   ||||||||||||||||||
N8        tggcgaggccagaggggcagcagaattggagttca
                  |||||||||||||||||
14-16     tggcgaggccagaggggcaTTACATCTGATGCAGT
                   ||||||||||||||||||
NX: 7-6   GGTAATTAGGAAGTGGCATTACATCTGATGCAGT
```

Figure 4. (**A**) Relationships of subclones 7a and 7b derived from cosmid Y61G2, and restriction maps of the junction fragment subclones of der(8) (9–2) and der(X) (14–16), and the normal X chromosome (7–6). The location and orientation of primers are shown. The restriction enzymes used were *Bgl*II (B), *Hin*dIII (H), *Sma*I (S) and *Xba*I (X). Filled and open bars indicate chromosome X and 8, respectively. (**B**) Nucleotide sequences of subclones from der(8) (9–2) and der(X) (14–16), normal chromosome X (NX: 7–6) and normal chromosome 8 (N8) spanning the translocation breakpoint. Chromosome X sequences are indicated in upper case and chromosome 8 sequences in lower case. The box indicates the 5′-GGCA-3′ sequence motif where the translocation occurred.

fusion gene between *SDC2* and *GRPR* was not possible from this translocation.

Exon trapping in the 8q22 region

Evidence of a gene other than *SDC2* in the region of the translocation was investigated by exon trapping (40) and partial genomic sequence analysis. Two cosmids (C2E and C9A) spanning the breakpoint were chosen; *SDC2* was present in one (C2E) but not in the other. Four exon-trapped products were purified and sequenced, one of which was produced from both cosmids. This product was purified and used as a probe to screen a testis cDNA library and a human multi-tissue Northern blot (Clontech). Both results were negative. The genomic sequence

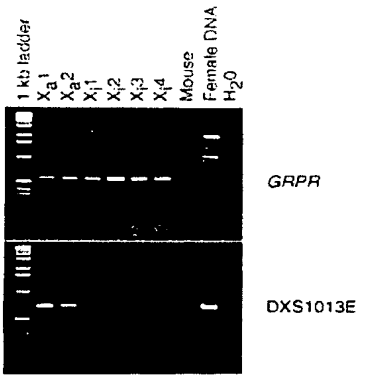


Figure 5. Escape from X inactivation of *GRPR*. The top figure demonstrates that *GRPR* is expressed from both the active and inactive X chromosomes. Primers used to detect expression are Ib1/B2 which amplify a product of 574 bp from cDNA, exon 2 and part of exon 3, and a larger product from DNA. *DXS1013E* expression is used as a control for a gene that is subject to inactivation. These primers amplify a product of ~700 bp. Lane 'mouse' refers to the parental mouse line (tsA1S9) and is a negative control for the amplification of murine products. The DNA lane corresponds to the female cell line GM7002. The lane on the far right is a H20 lane and molecular weight makers (1 kb ladder, Gibco-BRL) are in the far left lane.

containing this putative exon showed only one consensus splice site, indicating that it was probably a false positive. GRAIL, BLASTN and BLASTX analysis identified the three other exon-trapped products as Alu- and LINE-related sequences. No evidence was obtained for genes other than *SDC2* in this region by GRAIL, BLASTN and BLASTX searches of 41.7 kb of genomic sequence generated from *Hin*dIII subclones of cosmid C2E and C3H (Fig. 6).

Expression of the *SDC2* gene

The *SDC2* gene was expressed ubiquitously in human multi-tissue Northern blots (Clontech) (data not shown). To examine the expression level of *SDC2*, a Northern blot of total RNA isolated from the patient's lymphoblastoid cells was screened. The results indicated no expression of *SDC2* in either patient or control lymphoblastoid cells, but were positive for the expression in control fibroblast cells (data not shown). Thus expression level of *SDC2* in the patient remains unresolved since the patient's fibroblast cells were unavailable.

DISCUSSION

The translocation breakpoint of an X;8 translocation associated with multiple exostoses and autism was localized using a positional candidate cloning strategy. Sequencing confirmed that the breakpoint occurred within a 5′-GGCA-3′ sequence motif common to both X and 8 chromosomes, and to lie in intron 1 of the *GRPR* gene on the X chromosome, and <30 kb distal to the 3′ end of the *SDC2* gene on chromosome 8.

GRPR is a G-protein-coupled receptor for gastrin-releasing peptide (GRP), with 384 amino acids and seven transmembrane domains (41). GRP and neuromedin B (NMB) are the only two mammalian homologues of the amphibian peptide bombesin. Bombesin and GRP are identical in 90% of their C-terminal amino acids, which contains the full receptor-binding site. Like bombesin, GRP is involved in various neurobiological activities as a neurotransmitter, a paracrine hormone or a growth factor (42). It is found in the brain, gastrointestinal tract and lungs (42,43) although on multi-tissue Northern blot we detected only pancreas expression. This discrepancy may result from a higher sensitivity for detection of mRNA signals in specific regions on brain sections rather than mRNA extracted from whole brain tissue. The expression of *GRPR* mRNA in the brain has been studied by Wada *et al.* using *in situ* hybridization on rat brain tissues. The most intensive expression of *GRPR* mRNA was found throughout the limbic system (44). The limbic system is considered to be involved in emotionality, affect and attention.

As such, the biological functions of GRPR are consistent with a role in the pathophysiology of autism. Although we have found *GRPR* to clearly escape X-inactivation, the gene from the inactive X chromosome may nevertheless be expressed at only a fraction of the active allele (45). Such a low dosage might still be sufficient to cause autism in this patient. It is unknown whether inactivation patterns for a given gene are consistent throughout every tissue and cell within the body. This does not appear to be true for imprinted genes (46), so it is possible that *GRPR* could escape inactivation in test tissues such as fibroblasts and lymphoblasts (and other tissues) but could be subject to inactivation in the brain, resulting in no *GRPR* expression and the observed phenotype.

Confirmation of whether a defective *GRPR* gene causes autism should come from observing rearrangements in the *GRPR* gene in other autistic patients. Examination of unrelated autistic male patients for point mutations and altered splice sites in the *GRPR* gene and association studies using polymorphisms within and flanking the gene are clearly warranted. If a low dosage of the GRPR product would be sufficient to cause autism, examination of unrelated autistic male but also female patients for mutations in the *GRPR* gene should be undertaken.

The *SDC2* gene product is a member of a family of cell surface heparan sulfate proteoglycans that interact with adhesion molecules, growth factors and a variety of other effectors that support the shaping, maintenance and repair of an organism (47). One of its biological functions is related to aggregating cells in the formation of bone (48) and therefore is consistent with a possible role in multiple exostoses. However, we found it not to be disrupted by the translocation breakpoint and were unable to determine if its expression was altered in the patient. Translocations occurring in the 3′ end of the *PAX6* gene, as far away as ~85 or 125 kb, have been suggested to be disruptive to the gene by virtue of the position of the translocation (position effect) (49). The translocation reported here with respect to *SDC2* or other genes in 8q22.1 may be another example of this phenomenon.

MATERIALS AND METHODS

Fluorescence *in situ* hybridization (FISH)

YAC DNA was prepared in solution using 4.5 M guanidium hydrochloride extraction or GeneClean (Bioline 101 Inc.). Metaphase spreads were prepared from heparinized whole blood and also from an Epstein–Barr virus (EBV)-transformed lymphoblastoid cell line of the patient. Dehydrated slides in ethanol, after 3 min denaturation in 70% formamide in 2× SSC at 72°C, were hybridized *in situ* with YAC or cosmid clones

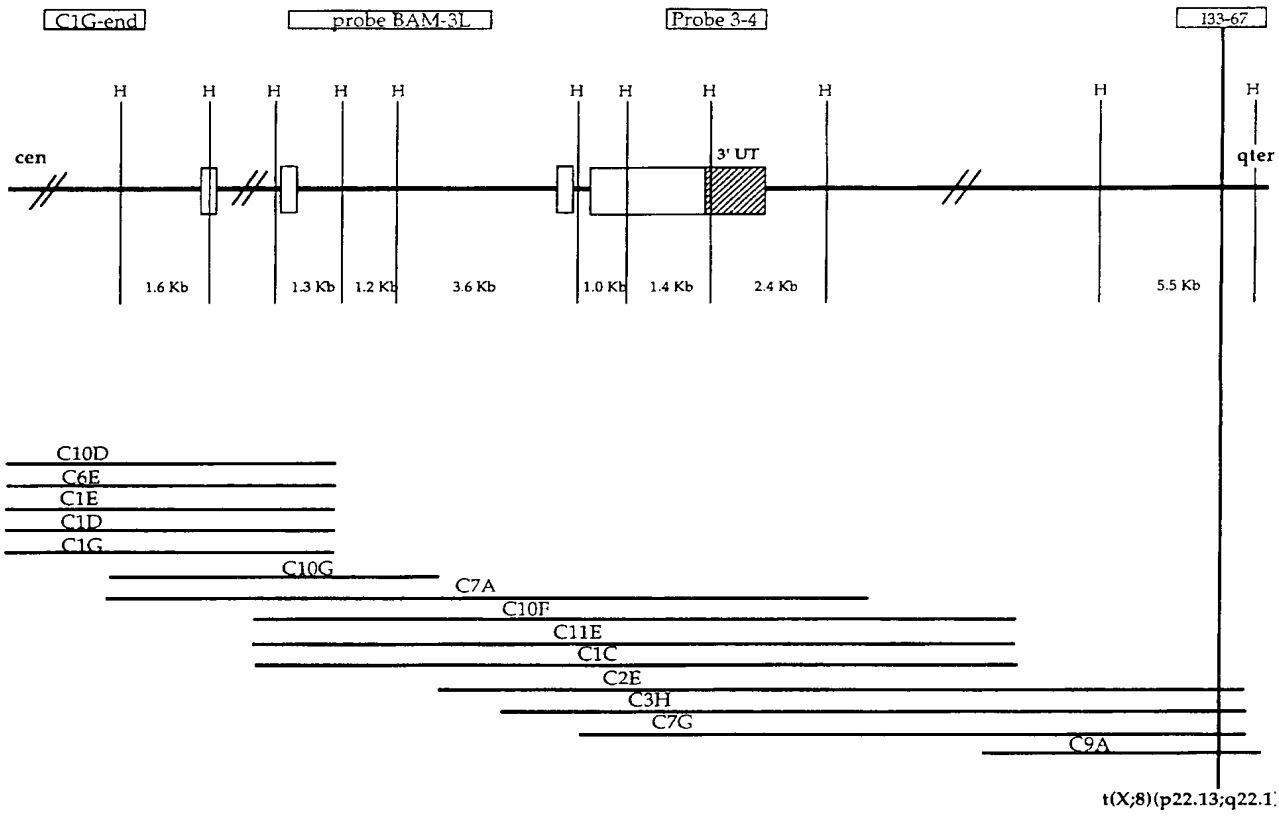


Figure 6. The location of the *SDC2* gene and ML's translocation breakpoint within the cosmid contig. The contig was constructed from the probes shown at the top of the figure.

labelled with biotin-14-dATP using nick translation (BioNick ™ Labeling System, Gibco, BRL). Interspersed repeat sequences of whole YAC/cosmid DNA was suppressed with Cot-1 DNA (Gibco, BRL) prior to hybridization. Hybridizations were performed as described in Buckle and Rack (50). The signals were detected by alternate layers of avidin–fluorescein isothiocyanate (FITC), biotinylated goat anti-avidin and avidin–FITC, the slides mounted in antifade (Vector Laboratories) containing propidium iodide and 4′,6′-diamidino-2-phenylindole (DAPI) and were captured as digital images with an MRC 600 scanning confocal microscope.

For the analysis of two probes with dual detection, one probe was labelled with biotin and the other with digoxigenin (DIG-11-dUTP; Boehringer-Mannheim). Biotin-labelled probes were detected with avidin–Texas red (Vector Laboratories), biotinylated goat anti-avidin and then avidin–Texas red. The DIG-labelled probes were detected by layers of anti-DIG (mouse) and anti-mouse FITC (goat) and counterstained with TO-PRO-3 (Molecular Probes) and DAPI. Weak Texas red signals occasionally were amplified further using a TSA-Indirect kit (Tyramide Signal Amplification, DuPont NEN). Two-colour signals were collected independently via an MRC 1024 scanning confocal microscope (Biorad) and collectively visualized using Lasersharp Software (Biorad).

Isolation of cosmids

Two methods were used for isolating cosmids: (i) screening the ICRF flow-sorted chromosome X cosmid library and (ii) constructing a cosmid sublibrary from YACs spanning the breakpoint. In the first method, YACs were excised from pulsed-field gel electrophoresis gels, radioactively labelled (T7 QuickPrime Kit, Pharmacia) and hybridized to the ICRF human flow-sorted X chromosome cosmid library (51,52). Cosmids were isolated and classified into either proximal and distal groups relative to the translocation breakpoint using FISH. In the second method, two cosmid libraries were constructed, one from YAC-60 (480 kb) and the other from YAC-61 (520 kb). These YACs were partially digested with 0.1 U of *Mbo*I (NEB) for 1 min at 37°C, resulting in fragments of 40–50 kb. One μg of dephosphorylated yeast DNA was ligated to 1 μg of *Bam*HI-digested SuperCos vector DNA (Stratagene) overnight at 16°C. *In vitro* packaging was performed using Gigapack II Gold packaging extracts (Stratagene). The cosmid libraries were plated on XL1 blue MR host strain and filter-lifted on Hybond N+ membranes. The filters were hybridized with radioactively labelled total human DNA, and cosmids with human inserts from each library were picked into microtitre plates and replicated to filters.

Construction of cosmid contig

Half of the cosmids isolated by the above procedures were digested with *Hin*dIII. A portion of the digested cosmid DNA was run on gels and blotted to filters and the remainder used as hybridization probes in the construction of a contig covering ~500 kb. Additional probes were prepared from six sequence-tagged sites (STSs) (33), four exon-specific products (amplified by primers from the published sequence of the *GRPR* Genbank accession number M73481) and 18 Alu-PCR products derived from the YAC clones.

Alu-PCR amplification

YACs spanning the translocation breakpoint were used as templates for Alu repeat-primed PCR using various combinations of two of either, Collins-5, Collins-3 (53), 2729 and 3144 (54). PCR was performed in a reaction buffer containing 10 mM Tris–HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 250 μM each dNTP, and YAC DNA at 500 ng/ml. An initial denaturation of 3 min at 94°C was followed by 30 cycles according to each set of primers. PCR products were analysed on ethidium bromide-stained 1% agarose gels and reloaded on to 1.3% SeaPlaque (F.M.C.) low melting agarose 1× TAE gel, which were excised, purified and used as probes.

Subcloning of cosmids

Cosmids spanning the translocation breakpoint and chosen for subcloning were digested with *Hin*dIII, ligated into *Hin*dIII-digested and dephosphorylated pBluescript, and transformed into competent XL1 Blue cells (Stratagene).

Hybridization of radioactively labelled probes to library filters

DNA was purified using Qiagen columns (Qiagen Inc.) and pre-annealed with placenta total human DNA (Sigma), vector DNA and DNA from an unrelated cosmid and then radioactively labelled by random hexanucleotide priming using [α-^{32}P]CTP (55) with T7 polymerase (QuickPrime kit, Pharmacia).

Cosmid and cDNA library filters were pre-hybridized overnight with 100 μg/ml denatured, sheared placenta total human DNA (Sigma) and then hybridized overnight with probe at a concentration of 1×10^6 c.p.m./ml. Hybridization was performed at 42°C in a buffer containing 50% formamide (56). The filter was washed in 1× SSC and 1% (w/v) SDS twice at 42°C and then twice at 65°C for 30 min, and was then exposed to Kodak XAR-5 film at –70°C using an intensifying screen.

Hybridization of non-radioactively labelled probes to Southern blots

DNA was non-isotopically labelled by the incorporation of DIG-11-dUTP via PCR or random-primed labelling methods. Oligonucleotides were labelled with terminal transferase adding by tailing with DIG-ddUTP with a mixture of nucleotides (Boehringer-Mannheim). Hybridization was performed in modified Church buffer (5% SDS, 0.5 M sodium phosphate) (57). The blots were washed twice in 2× SSC and 0.1% (w/v) SDS at room temperature for 10 min and then once in 0.1× SSC and 0.1% (w/v) SDS at 65°C for 20 min and mixed with anti-DIG antibody. Results were detected by applying Attophos (EUROPA) with a sponge and visualized by photographing under long wave UV (365 nm).

Preparation of patient's Northern blot

Total cellular RNA was extracted from an EBV-transformed lymphoblastoid cell line using TRIzol (Gibco, BRL). Thirty μg of RNA was separated through a 1% agarose gel containing 6% formaldehyde and transferred to a Hybond N+ (Amersham) filter.

Construction of genomic phage library from ML

The size of ML's *Bgl*II-digested junction fragments (5.5 kb) was within 1 kb of the normal X chromosome *Bgl*II fragment (4.6 kb). As this may facilitate equal representation after cloning into phage, *Bgl*II was used to digest ML's genomic DNA for construction of the phage library. A 3-genome fold library was plated using the XL1 Blue MRA host strain, transferred on to Hybond N^+ (Amersham), and screened using two single-copy plasmid DNA subclones (7a and 7b) isolated from the cosmid Y61G2 which detected junction fragments.

GRPR gene expression on active and inactive X chromosomes

Cell lines were grown in α-MEM media supplemented with 10% fetal calf serum, penicillin and streptomycin, and L-glutamine. Hybrids were maintained under the appropriate selection: AHA-11aB1 was grown in HAT at 37°C, t60-12 was grown in HAT at 39°C. Cells were harvested from a single t175 flask just prior to confluency for RNA isolation (36). RNA was extracted from cell lines with RNAzol (Cinna/Biotecx) using two sequential extractions of the RNA to eliminate DNA contamination completely. Reverse transcription reactions were done with and without reverse transcriptase to control for DNA contamination within the RNA sample. Approximately 5 μg of total RNA was reverse transcribed in PCR buffer (Gibco BRL) using random hexamers to prime the RNA strand. The resulting cDNA samples were diluted 1:2 and 1 μl was subjected to PCR with gene-specific primers. PCR conditions consisted of an initial heat denaturation at 95°C for 2 min followed by 35 cycles of 94°C for 15 s; 55°C for 15 s; 72° for 40 s in a 50 μl reaction with 1.5 mM $MgCl_2$. PCR products were run out on 2.0% agarose gels, visualized with ethidium bromide and photographed.

Exon trapping

Cosmids C9A and C7B, which span the site of the translocation breakpoint, were digested with *Bgl*II and *Bam*HI and ligated into the *Bam*HI site of the pSPL3 exon-trapping vector (40). pSPL3 clones containing cosmid inserts were Qiagen-prepared (Qiagen Inc.) and transfected by electroporation into COS-7 cells. After 72 h of growth, the cells were harvested and RNA was extracted using TRIzol (Gibco, BRL) kit. The first strand of cDNA was synthesized using the SA2 primer and reverse transcriptase (NEB). The first round of PCR amplification was performed on the cDNA using the SA2 and SD6 primers. The PCR products larger than the 110 bp vector product were separated in 1.3% SeaPlaque (F.M.C.) low melting agarose gel, excised and purified using the QiaQuick PCR Purification kit (Qiagen Inc.). The resulting products were PCR-amplified again using the nested primers SA4 and SD2. The final PCR products from the second round of PCR amplification were sized on gels, excised, purified

and then ligated into the pAMP1 vector (Clone Amp System, BRL) and sequenced.

DNA sequencing and generation of primers and STSs

Cycle sequencing was performed on double-stranded pBluescript purified using the Magic miniprep kit (Promega) or PCR fragments purified using a PCR Preps DNA Purification Resin (Promega) with *Taq* polymerase, fluorescent dideoxy chain terminators and Applied Biosystems buffer. The reaction was undertaken in a thermocycler (Perkin Elmer) for 25 cycles of 96°C for 30 s, 50°C for 15 s and 60°C for 4 min. Samples purified with phenol–chloroform extraction were pelleted with ethanol and subjected to electrophoresis on an ABI 373A DNA Sequencer. The sequence data were analysed using Factura and Sequencing Navigator software (Applied Biosystems). The edited sequences were analysed by GRAIL (58), which uses a multi-sensor neural network approach to find genes in DNA sequences, and also analysed by BLASTX (59), which translates nucleotide sequences in all six reading frames and searches a non-redundant protein database. Nucleotide homologies were also searched using BLASTN. All BLAST searches were done by the NCBI server and GRAIL analysis by the ORNL server.

ACKNOWLEDGEMENTS

We are grateful to A. P. Walker, I. W. Craig and A. Hovnanian for stimulating discussions, K. A. Rack, V. J. Buckle and E. R. Levy for help in preparing metaphase spreads for FISH analysis, C. Philippe for help with exon trapping, I. Millwood, C. Stark, E. Maestrini and A. Nemeth for critical comments on the manuscript, and I. Randall for help with references. This work was funded by the Imperial Cancer Research Fund, the Wellcome Trust and the National Institutes of Health. A.P.M. is a Wellcome Trust Principal Research Fellow.

REFERENCES

1. Mandel, J.-L., Monaco, A.P., Nelson, D.L., Schlessinger, D. and Willard, H.F. (1992) Genome analysis and the human X chromosome. *Science*, **258**, 103–109.
2. Bolton, P., Powell, J., Rutter, M., Buckle, V., Yates, J.R.W., Ishikawa-Brush, Y. and Monaco, A.P. (1995) Autism, mental retardation, multiple exostoses and short stature in a female with 46,X,t(X;8)(p22.13;q22.1). *Psychiatr. Genet.*, **5**, 51–55.
3. Le Couteur, A., Rutter, M., Lord, C., Rios, P., Robertson, S., Holgrafer, M. and McLennan, J. (1989) Autism diagnostic interview: a standardized investigator-based instrument. *J. Autism Dev. Disord.*, **19**, 363–387.
4. Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L. and Schopler, E. (1989) Autism diagnostic observation schedule: a standardized observation of communicative and social behaviour. *J. Autism Dev. Disord.*, **19**, 185–212.
5. McKusick, V.A. (1994) *Mendelian Inheritance in Man*. The Johns Hopkins University Press, Baltimore, MD.
6. Rutter, M. (1970) Autistic children: infancy to adulthood. *Semin. Psychiatr.*, **2**, 435–450.
7. Smalley, S.L., Asarnow, R.F. and Spence, A. (1988) Autism and genetics: a decade of research. *Arch. Gen. Psychiatr.*, **45**, 953–961.
8. Folstein, S. and Rutter, M. (1977) Infantile autism: a genetic study of 21 twin pairs. *J. Child Psychol. Psychiatr.*, **18**, 297–321.
9. Steffenburg, S., Gillberg, C., Hellgren, L., Andersson, L., Gillberg, I., Jakobsson, G. and Bohman, M. (1989) A twin study of autism in Denmark, Finland, Iceland, Norway and Sweden. *J. Child Psychol. Psychiatr.*, **30**, 405–416.
10. Bolton, P. and Rutter, M. (1990) Genetic influences in autism. *Int. Rev. Psychiatr.*, **2**, 67–80.
11. Risch, N. (1990) Linkage strategies for genetically complex traits I. Multilocus models. *Am. J. Hum. Genet.*, **46**, 222–228.
12. Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E. and Rutter, M. (1995) Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol. Med.*, **25**, 63–77.
13. Blomquist, H.K., Bohman, M., Edvinsson, S.O., Gillberg, C., Gustavson, K.-H., Holmgren, G. and Wahlstrom, J. (1985) Frequency of the fragile X syndrome in infantile autism: a Swedish multicenter study. *Clin. Genet.*, **27**, 113–117.
14. Bailey, A. (1993) The biology of autism. *Psychol. Med.*, **23**, 7–11.
15. Hallmayer, J., Spiker, D., Lotspeich, L., McMahon, W.M., Petersen, P.B., Nicholas, P., Pingree, C. and Ciaranello, R.D. (1994) Male-to-male transmission in extended pedigrees with multiple cases of autism. *Am. J. Med. Genet.*, **67**, 13–18.
16. Hallmayer, J., Hebert, J.M., Spiker, D., Lotspeich, L., McMahon, W.M., Petersen, P.B., Nicholas, P., Pingree, C., Lin, A.A., CavalliSforza, L.L., Risch, N. and Ciaranello, R.D. (1996) Autism and the X chromosome: multipoint sib-pair analysis. *Arch. Gen. Psychiatr.*, **53**, 985–989.
17. Eden, M.S., Thelin, J.W., Michalski, K. and Mitchel, J.A. (1985) Partial trisomy 6p and partial monosomy 9p from a *de novo* translocation 46,XY,-9,+der(9)t(6:9)(p21:p24). *Clinic. Genet.*, **28**, 375–384.
18. Mariner, R., Jackson III, A.W., Levitas, A., Hagerman, R.J., Braden, P.M., McBogg, P.M., Smith, A.C.M. and Berry, R. (1986) Autism, mental retardation, and chromosomal abnormalities. *J. Autism Dev. Disord.*, **16**, 425–440.
19. Burd, L., Martsolf, J.T., Kerbeshian, J. and Jalal, S.M. (1988) Partial 6p trisomy associated with infantile autism. *Clinic. Genet.*, **33**, 356–359.
20. Gillberg, C. and Coleman, M. (1992) *Clinics in Developmental Medicine. 2nd edn., The Biology of the Autistic Syndrome*. Mac Keith Press, London.
21. Tully, E.M., Mellon, C.D. and McMahon, W.M. (1993) Autism associated with 9p-deletion syndrome. *Biol. Psychiatr.*, **33**, 37A–162A.
22. Ghaziuddin, M., Sheldon, S., Tsai, L.Y. and Alessi, N. (1993) Abnormalities of chromosome 18 in a girl with mental retardation and autistic disorder. *J. Intellectual Disability Res.*, **37**, 313–317.
23. Vostanis, P., Harrington, R., Prendergast, M. and Farndon, P. (1994) Case reports of autism with interstitial deletion of chromosome 17 (p11.2 p11.2) and monosomy of chromosome 5 (5pter→5p15.3). *Psychiatr. Genet.*, **4**, 109–111.
24. Rao, P.N., Klinepeter, K., Stewart, W., Hayworth, R., Grubs, R. and Pettenati, M.J. (1994) Molecular cytogenetic analysis of a duplication Xp in a male: further delineation of a possible sex influencing region on the X chromosome. *Hum. Genet.*, **94**, 149–153.
25. Schmale, G.A., Conrad, E.U. and Raskid, W.H. (1994) The natural history of hereditary multiple exostoses. *J. Bone Joint Surg. Am.*, **76**, 986–992.
26. Cook, A., Raskind, W., Blanton, S., H. Pauli, R.M., Gregg, R.G., Francomano, C.A., Puffenberger, E., Conrad, E.U., Schmale, G., Schellenberg, G., Wijsman, E., Hecht, J.T., Wells, D. and Wagner, M.J. (1993) Genetic heterogeneity in families with hereditary multiple exostoses. *Am. J. Hum. Genet.*, **53**, 71–79.
27. Ahn, J., Ludecke, H.-J., Lindow, S., Horton, W.A., Lee, B., Wagner, M.J., Horsthemke, B. and Wells, D.E. (1995) Cloning of the putative tumour suppressor gene for hereditary multiple exostoses (EXT1). *Nature Genet.*, **11**, 137–143.
28. Wuyts, W., Van Hul, W., Wauters, J., Nemtsova, M., Reyniers, E., Van Hul, E., De Boulle, K., de Vries, B.B.A., Hendrickx, J., Herrygers, I., Bossuyt, P., Balemans, W., Fransen, E., Vits, L., Coucke, P., Nowak, N.J., Shows, T.B., Mallet, L., van den Ouweland, A.M.W., McGaughran, J., Halley, D.J.J. and Willems, P.J. (1996) Positional cloning of a gene involved in hereditary multiple exostoses. *Hum. Mol. Genet.*, **5**, 1547–1557.
29. Stickens, D., Clines, G., Burbee, D., Ramos, P., Thomas, S., Hogue, D., Hecht, J.T., Lovett, M. and Evans, G.A. (1996) The EXT2 multiple exostoses gene defines a family of putative tumour suppressor genes. *Nature Genet.*, **14**, 25–32.
30. Hecht, J.T., Hogue, D., Strong, L.C., Hansen, M.F., Blanton, S.H. and Wagner, M. (1995) Hereditary multiple exostosis and chondrosarcoma: linkage to chromosome 11 and loss of heterozygosity for EXT-linked markers on chromosomes 11 and 8. *Am. J. Hum. Genet.*, **56**, 1125–1131.
31. Yoshiura, K., Inazawa, J., Koyama, K., Nakamura, Y. and Niikawa, N. (1994) Mapping of the 8q23 translocation breakpoint of t(8;13) observed in a patient with multiple exostoses. *Genes, Chromosomes Cancer*, **9**, 57–61.

32. Larin, Z., Monaco, A.P. and Lehrach, H. (1991) Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. *Proc. Natl Acad. Sci. USA*, **88**, 4123–4127.
33. Francis, F., Benham, F., See, C.G., Fox, M., Ishikawa-Brush, Y., Monaco, A.P., Weiss, B., Rappold, G., Hamvas, R.M.J. and Lehrach, H. (1994) Identification of YAC and cosmid clones encompassing the ZEX-POLA region using irradiation hybrid cell lines. *Genomics*, **20**, 75–85.
34. Alitalo, T., Francis, F., Kere, J., Lehrach, H., Schlessinger, D. and Willard, H.F. (1995) A 6-Mb YAC contig in Xp22.1–Xp22.2 spanning the DXS69E, XE59, GLRA2, PIGA, GRPR, CALB3 and PHKA2 genes. *Genomics*, **25**, 691–700.
35. Corjay, M.H., Dobrzanski, D.J., Way, J.M., Viallet, J., Shapira, H., Worland, P., Sausville, E.A. and Battey, J.F. (1991) Two distinct bombesin receptor subtypes are expressed and functional in human lung carcinoma cells. *J. Biol. Chem.*, **266**, 18771–18779.
36. Brown, C.J., Flenniken, A.M., Williams, B.R.G. and Willard, H.F. (1990) X chromosome inactivation of the human TIMP gene. *Nucleic Acids Res.*, **18**, 4191–4195.
37. Willard, H.F., Brown, C.J., Carrel, L., Hendrich, B. and Miller, A.P. (1993) Epigenetic and chromosomal control of gene expression: molecular and genetic analysis of X chromosome inactivation. *Cold Spring Harbor Symp. Quant. Biol.*, **58**, 315–322.
38. Miller, A.P., Gustashaw, K., Wolff, D.J., Rider, S.H., Monaco, A.P., Eble, B., Schlessinger, D., Gorski, J.L., van Ommen, G.J., Weissenbach, J. and Willard, H.F. (1995) Three genes that escape X chromosome inactivation are clustered with a 6 Mb YAC contig and STS map in Xp11.21–p11.22. *Hum. Mol. Genet.*, **4**, 731–739.
39. Marynen, P., Zhang, J., Cassiman, J.-J., Van den Berghe, H. and David, G. (1989) Partial primary structure of the 48- and 90-kilodalton core proteins of cell surface-associated heparan sulfate proteoglycans of lung fibroblasts. *J. Biol. Chem.*, **264**, 7017–7024.
40. Buckler, A.J., Chang, D.D., Graw, S.L., Brook, J.D., Haber, D.A., Sharp, P.A. and Housman, D.E. (1991) Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. *Proc. Natl Acad. Sci. USA*, **88**, 4005–4009.
41. Dohlman, H.G., Caron, M.G. and Lefkowitz, R.J. (1987) A family of receptors coupled to guanine-nucleotide regulatory proteins. *Biochemistry*, **26**, 2657–2663.
42. Spindel, E.R., Zilberberg, M.D., Habener, J.F. and Chin, W.W. (1986) Two prohormones for gastrin-releasing peptide are encoded by two mRNAs differing by 19 nucleotides. *Proc. Natl Acad. Sci. USA*, **83**, 19–23.
43. Wang, D. and Cutz, E. (1994) Simultaneous detection of messenger ribonucleic acids for bombesin/gastrin-releasing peptide and its receptor in rat brain by nonradiolabeled double *in situ* hybridization. *Lab. Invest.*, **70**, 775–780.
44. Wada, E., Way, J., Shipira, H., Kusano, K., Lebacq-Verheyden, A., Coy, D., Jensen, R. and Battery, J. (1991) cDNA cloning, characterization and brain region-specific expression of a neuromedin-β-preferring bombesin receptor. *Neuron*, **6**, 421–430.
45. Migeon, B.R., Shapiro, L.J., Norum, R.A., Mohandas, T., Axelman, J. and Dabora, R.L. (1982) Differential expression of steroid sulphatase locus on active and inactive human X chromosome. *Nature*, **299**, 838–840.
46. Lee, M.P., Hu, R.-J., Johnson, L.A. and Feinberg, A.P. (1997) Human *KVLQT1* gene shows tissue-specific imprinting and encompasses Beckwith–Wiedemann syndrome chromosomal rearrangements. *Nature Genet.*, **15**, 181–185.
47. David, G., Bai, X.M., Van der Schueren, B., Marynen, P., Cassiman, J.-J. and Van den Berghe, H. (1993) Spatial and temporal changes in the expression of fibroglycan (Syndecan-2) during mouse embryonic development. *Development*, **119**, 841–854.
48. Solursh, M., Reiter, R.S., Jensen, K.L., Kato, M. and Bernfield, M. (1990) Transient expression of a cell surface heparan sulfate proteoglycan (Syndecan) during limb development. *Dev. Biol.*, **140**, 83–92.
49. Fantes, J., Redeker, B., Breen, M., Boyle, S., Brown, J., Fletcher, J., Jones, S., Bickmore, W., Fukushima, Y., Mannens, M., Danes, S., van Heyningen, V. and Hanson, I. (1995) Aniridia-associated cytogenetic rearrangments suggest that a position effect may cause the mutant phenotype. *Hum. Mol. Genet.*, **4**, 415–422.
50. Buckle, V.J. and Rack, K.A. (1993) Fluorescent *in situ* hybridization. In Davies, K.E. (ed.) *Human Genetic Disease Analysis—A Practical Approach*. IRL Press, Oxford, pp. 59–82.
51. Nizetic, D., Zehetner, G., Monaco, A.P., Gellen, L., Young, B.D. and Lehrach, H. (1991) Construction, arraying and high density screening of large insert cosmid libraries of the human chromosomes X and 21. *Proc. Natl Acad. Sci. USA*, **88**, 3233–3237.
52. Baxendale, S., Bates, G.P., MacDonald, M.E., Gusella, J.F. and Lehrach, H. (1991) The direct screening of cosmid libraries with YAC clones. *Nucleic Acids Res.*, **19**, 6651.
53. Tagle, D. and Collins, S. (1992) An optimized Alu-PCR primer pair for human specific amplification of YACs somatic cell hybrid. *Hum. Mol. Genet.*, **1**, 121–122.
54. Monaco, A.P., Lam, V.M.S., Zehetner, G., Lennon, G.G., Douglas, C., Nizetic, D., Goodfellow, P.N. and Lehrach, H. (1991) Mapping irradiation hybrids to cosmid and yeast artificial chromosome libraries by direct hybridization of Alu-PCR products. *Nucleic Acids Res.*, **19**, 3315–3318.
55. Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.*, **132**, 6–13.
56. Monaco, A.P., Bertleson, C.J., Middlesworth, W., Colletti, C.A., Fischbeck, K.H., Bartlett, R., Pericack-Vance, M.A., Roses, A.D., Aldridge, J. and Kunkel, L.M. (1985) Detection of deletions spanning the Duchenne muscular dystrophy locus with a DNA segment tightly linked to the disease. *Nature*, **316**, 842–845.
57. Church, G.M. and Gilbert, W. (1984) Genomic sequencing. *Proc. Natl Acad. Sci. USA*, **81**, 1991–1995.
58. Uberbacher, E.E. and Mural, R.J. (1991) Locating protein-coding regions in human DNA sequences by a multiple sensor-neural network approach. *Proc. Natl Acad. Sci. USA*, **88**, 11261–11265.
59. Gish, W. and States, D. (1993) Identification of protein coding regions by database similarity search. *Nature Genet.*, **3**, 266–272.